

WO 00/52042

PCT/EP00/01468

1/20

Figure 1 : Alignment of the BASB081 polynucleotide sequences.

Identity to SeqID No:1 is indicated by a dot. Gap is indicated by a dash.

|          |                                |     |     |     |
|----------|--------------------------------|-----|-----|-----|
|          | *                              | 20  | *   |     |
| Seqid1 : | ATGTCAAAGCCCGTTTTGTTTGCAAATCGC | :   |     | 30  |
| Seqid3 : | -----                          | :   |     | -   |
|          |                                |     |     |     |
|          | 40                             | *   | 60  |     |
| Seqid1 : | AGTTTTATGCCTGTCGCATTGGCGGCTTAT | :   |     | 60  |
| Seqid3 : | -----                          | :   |     | -   |
|          |                                |     |     |     |
|          | *                              | 80  | *   |     |
| Seqid1 : | TTGCCTTTGATGACATCGCAAGCATTGGCA | :   |     | 90  |
| Seqid3 : | -----                          | :   |     | -   |
|          |                                |     |     |     |
|          | 100                            | *   | 120 |     |
| Seqid1 : | CAACAAAATAACCCTGCAAACATCATCAAT | :   |     | 120 |
| Seqid3 : | .....                          | :   |     | 30  |
|          |                                |     |     |     |
|          | *                              | 140 | *   |     |
| Seqid1 : | CATGTACCCGCTCATGACACCGCCATCAAT | :   |     | 150 |
| Seqid3 : | .....                          | :   |     | 60  |
|          |                                |     |     |     |
|          | 160                            | *   | 180 |     |
| Seqid1 : | CAAGCAAAGGCAGGCAATCCGCCTGTTTTG | :   |     | 180 |
| Seqid3 : | .....                          | :   |     | 90  |

WO 00/52042

PCT/EP00/01468

2/20

|        |   |                                |   |     |
|--------|---|--------------------------------|---|-----|
|        | * | 200                            | * |     |
| Seqid1 | : | CTAACACCTGAGCAGATACAAGCACGCCTT | : | 210 |
| Seqid3 | : | .....                          | : | 120 |

|        |     |                                |     |     |
|--------|-----|--------------------------------|-----|-----|
|        | 220 | *                              | 240 |     |
| Seqid1 | :   | AATGCTGCTGGACTGAATGCTAAGCCCCAA | :   | 240 |
| Seqid3 | :   | .....                          | :   | 150 |

|        |   |                                |   |     |
|--------|---|--------------------------------|---|-----|
|        | * | 260                            | * |     |
| Seqid1 | : | TCACAAGCTTTGGATGTTGTCAATTTTGAT | : | 270 |
| Seqid3 | : | .....                          | : | 180 |

|        |     |                                |     |     |
|--------|-----|--------------------------------|-----|-----|
|        | 280 | *                              | 300 |     |
| Seqid1 | :   | GATCAATCGCCGATATCTCGTATCGGTGAG | :   | 300 |
| Seqid3 | :   | .....                          | :   | 210 |

|        |   |                                |   |     |
|--------|---|--------------------------------|---|-----|
|        | * | 320                            | * |     |
| Seqid1 | : | CAATCACCCCCTTTGGGTTTGGATATGTCG | : | 330 |
| Seqid3 | : | .....                          | : | 240 |

|        |     |                                |     |     |
|--------|-----|--------------------------------|-----|-----|
|        | 340 | *                              | 360 |     |
| Seqid1 | :   | GTCATCGAAGAAACCACACCGCTAAGCTTG | :   | 360 |
| Seqid3 | :   | .....                          | :   | 270 |

|        |   |                                |   |     |
|--------|---|--------------------------------|---|-----|
|        | * | 380                            | * |     |
| Seqid1 | : | GAGGAATTATTTGCTCAAGAATCTACTGAG | : | 390 |
| Seqid3 | : | .....                          | : | 300 |

WO 00/52042

PCT/EP00/01468

3/20

400 \* 420  
 Seqid1 : ATGGGAATCAATCCAAATGATTATATTCCA : 420  
 Seqid3 : ..... : 330

\* 440 \*  
 Seqid1 : GAATATCAAGGCGAGCAACCTAATAGTGAG : 450  
 Seqid3 : ..... : 360

460 \* 480  
 Seqid1 : GTGGTTGTACCACCGACATTAGAACCTGAA : 480  
 Seqid3 : ..... : 390

\* 500 \*  
 Seqid1 : AAACCAGGTTTGATCAAGCGTCTTTATGCA : 510  
 Seqid3 : ..... : 420

520 \* 540  
 Seqid1 : CGCCTATTTAATGATGGTGTCAATAAGGTG : 540  
 Seqid3 : ..... : 450

\* 560 \*  
 Seqid1 : CCTAGGCTTAAGGCAAAATTTTATCAATCA : 570  
 Seqid3 : ..... : 480

580 \* 600  
 Seqid1 : TCGCAATCAGGCGAAACCAGTGCGATTGGG : 600  
 Seqid3 : ..... : 510

WO 00/52042

PCT/EP00/01468

4/20

\* 620 \*

Seqid1 : TCATCGCATCAAAAAACAGAGCCTTATGCA : 630

Seqid3 : ..... : 540

640 \* 660

Seqid1 : AATATCAAAGCAGCACTTGAAGACATCACC : 660

Seqid3 : ..... : 570

\* 680 \*

Seqid1 : CAAGAGTCAGCGATGGATTTGAATGGCTCT : 690

Seqid3 : ..... : 600

700 \* 720

Seqid1 : ATCCACGCCTAAGGCAAAGTCTTTGGTG : 720

Seqid3 : ..... : 630

\* 740 \*

Seqid1 : GCAGCGCGTGCTGTCGGTTATTATGATATT : 750

Seqid3 : ..... : 660

760 \* 780

Seqid1 : GATTTATCAATCATAAGAAATAGCATCGGA : 780

Seqid3 : ..... : 690

\* 800 \*

Seqid1 : GAGGTGGATGTCATCATCCATGATTTAGGT : 810

Seqid3 : ..... : 720

WO 00/52042

PCT/EP00/01468

5/20

820 \* 840  
 Seqid1 : GAACCTGTTTATATTGATTATCGAGCGGTG : 840  
 Seqid3 : ..... : 750

\* 860 \*  
 Seqid1 : GAGGTACGAGGTGAAGGTGCTGATGATAAA : 870  
 Seqid3 : ..... : 780

880 \* 900  
 Seqid1 : GCATTTACTACCGTGGCGGATGAGGTGCCA : 900  
 Seqid3 : ..... : 810

\* 920 \*  
 Seqid1 : TTGCTGATCGGCGATGTCTTTCATCATGGC : 930  
 Seqid3 : .....C... : 840

940 \* 960  
 Seqid1 : AAGTACGAAACCAAAAAAATCTCATCGAA : 960  
 Seqid3 : ..... : 870

\* 980 \*  
 Seqid1 : AATGCCAGTGCTGAACATGGATATTTTGAT : 990  
 Seqid3 : ..... : 900

1000 \* 1020  
 Seqid1 : GGGCGTTGGCTGGATCGTTCAGTTGATGTA : 1020  
 Seqid3 : ..... : 930

WO 00/52042

PCT/EP00/01468

6/20

```
          *          1040          *
Seqid1  : ATTTTGCCAGATAATACCGCTGATGTCAGC : 1050
Seqid3  : ..... : 960
```

```
          1060          *          1080
Seqid1  : TTAATTTATGATACAGGTACGCAGTATCGC : 1080
Seqid3  : ..... : 990
```

```
          *          1100          *
Seqid1  : TTTGATGAGGTGGTATTTTTTACCATTGAT : 1110
Seqid3  : .....A..... : 1020
```

```
          1120          *          1140
Seqid1  : CCTAAAACCAATCAATTGACAACCGATCCA : 1140
Seqid3  : ..... : 1050
```

```
          *          1160          *
Seqid1  : GATAAGCTGCCAGTTAAACGAGAATTACTT : 1170
Seqid3  : ..... : 1080
```

```
          1180          *          1200
Seqid1  : GAGCAGTTACTCACCGTTAACATGGGAGAG : 1200
Seqid3  : ..... : 1110
```

```
          *          1220          *
Seqid1  : GCTTACAATTTACAGGCGGTGCGTGCACTT : 1230
Seqid3  : ..... : 1140
```

WO 00/52042

PCT/EP00/01468

7/20

1240 \* 1260  
Seqid1 : TCAAATGATTTGATTGCCACACGGTATTTT : 1260  
Seqid3 : ..... : 1170

\* 1280 \*  
Seqid1 : AATATGGTGAATACCGAGATTGTCTTTCCA : 1290  
Seqid3 : ..... : 1200

1300 \* 1320  
Seqid1 : GAGCGTGAACAGATCCAAAACGACCAAGTG : 1320  
Seqid3 : ..... : 1230

\* 1340 \*  
Seqid1 : AGCTTTGAGCAGTCTTCAAGTAGCCGTACT : 1350  
Seqid3 : ..... : 1260

1360 \* 1380  
Seqid1 : GAACCAGCACAAGTTGATGAAAGCACACTT : 1380  
Seqid3 : ..... : 1290

\* 1400 \*  
Seqid1 : GAACCTGTCATTGAAACCGTTGAGCTAACG : 1410  
Seqid3 : ..... : 1320

1420 \* 1440  
Seqid1 : GATGGGATATTAATGGATATTTGCCCCATC : 1440  
Seqid3 : ..... : 1350

WO 00/52042

PCT/EP00/01468

8/20

\* 1460 \*

Seqid1 : GAATTTAGTGCATCTAATCTGATTCAAGAC : 1470

Seqid3 : ..... : 1380

1480 \* 1500

Seqid1 : AAGCTAAATTTGGTGGCTGCCAAGGCTCGC : 1500

Seqid3 : ..... : 1410

\* 1520 \*

Seqid1 : CATTTATATGACATGCCTGATGATAGGGTG : 1530

Seqid3 : ..... : 1440

1540 \* 1560

Seqid1 : CTTGCCATCAATCATGATGATGGCGTAAAT : 1560

Seqid3 : ..... : 1470

\* 1580 \*

Seqid1 : CGCTCTATTTTGGGCAGAATCAGCGATGCC : 1590

Seqid3 : ..... : 1500

1600 \* 1620

Seqid1 : GTATCTGCCGTTGCACGTGCTATTTTACCT : 1620

Seqid3 : ..... : 1530

\* 1640 \*

Seqid1 : GATGAATCTGAAAATGAGGTAATAGATTTG : 1650

Seqid3 : ..... : 1560



WO 00/52042

PCT/EP00/01468

9/20

1660 \* 1680  
Seqid1 : CCCGAGCGTACCGCATTGGCTAATCGCAAG : 1680  
Seqid3 : ..... : 1590

\* 1700 \*  
Seqid1 : ACCCCTGCTGATGTCTATCAAAGTAAAAAA : 1710  
Seqid3 : ..... : 1620

1720 \* 1740  
Seqid1 : GTGCCGCTATATGTCTTTGTGGCGAGTGAT : 1740  
Seqid3 : ..... : 1650

\* 1760 \*  
Seqid1 : AAACCACGAGATGGTCAAATTGGTTTGGGC : 1770  
Seqid3 : .....C..... : 1680

1780 \* 1800  
Seqid1 : TGGGGATCGGACACAGGTACCCGCCTAGTC : 1800  
Seqid3 : ..... : 1710

\* 1820 \*  
Seqid1 : ACAAAATTTGAGCATAATTTGATTAATCGT : 1830  
Seqid3 : ..... : 1740

1840 \* 1860  
Seqid1 : GATGGCTATCAAGCAGGCGCTGAGCTAAGA : 1860  
Seqid3 : ..... : 1770

WO 00/52042

PCT/EP00/01468

10/20

\* 1880 \*

Seqid1 : CTGTCTGAGGATAAAAAAGGGGTCAAGTTA : 1890

Seqid3 : ..... : 1800

1900 \* 1920

Seqid1 : TATGCCACCAAACCGCTTAGCCACCCTCTA : 1920

Seqid3 : ..... : 1830

\* 1940 \*

Seqid1 : AATGATCAGCTAAGAGCAACTTTGGGTAT : 1950

Seqid3 : ..... : 1860

1960 \* 1980

Seqid1 : CAACAAGAAGTTTTTGGTCACTCTACCAAT : 1980

Seqid3 : ..... : 1890

\* 2000 \*

Seqid1 : GGTTTTGATTTATCCACACGCACCCTAGAG : 2010

Seqid3 : ..... : 1920

2020 \* 2040

Seqid1 : CATGAGATTAGCCGCAGTATTATCCAAAAT : 2040

Seqid3 : ..... : 1950

\* 2060 \*

Seqid1 : GGTGGCTGGAATCGTACTTATTCATTGCGT : 2070

Seqid3 : ..... : 1980

WO 00/52042

PCT/EP00/01468

11/20

2080 \* 2100  
Seqid1 : TATCGTCTTGATAAGCTTAAAACCCAAGCA : 2100  
Seqid3 : ..... : 2010

\* 2120 \*  
Seqid1 : CCCCCTGAAACATGGCAGGATTACCAGTG : 2130  
Seqid3 : ..... : 2040

2140 \* 2160  
Seqid1 : GATTTTGTCAATGGTAAGCCAAGCCAAGAG : 2160  
Seqid3 : ..... : 2070

\* 2180 \*  
Seqid1 : GCGTTATTGGCAGGTGTTGCTGTGCATAAA : 2190  
Seqid3 : ..... : 2100

2200 \* 2220  
Seqid1 : ACGGTTGCAGATAATTTGGTTAATCCGATG : 2220  
Seqid3 : ..... : 2130

\* 2240 \*  
Seqid1 : CGTGGCTATCGTCAGCGATATTCTTTAGAG : 2250  
Seqid3 : ..... : 2160

2260 \* 2280  
Seqid1 : GTTGGCTCAAGCGGTTTGGTATCGGATGCT : 2280  
Seqid3 : ..... : 2190

WO 00/52042

PCT/EP00/01468

12/20

\* 2300 \*

Seqid1 : AATATGGCTATTGCTCGAGCTGGTATTAGT : 2310

Seqid3 : ..... : 2220

2320 \* 2340

Seqid1 : GGCGTGTATAGTTTTGGGGATAATGCTTAT : 2340

Seqid3 : ..... : 2250

\* 2360 \*

Seqid1 : GGCAGCAATCGTGCCCATCAGATGACTGGT : 2370

Seqid3 : ..... : 2280

2380 \* 2400

Seqid1 : GGCATACAAGCAGGATACATTTGGTCGGAT : 2400

Seqid3 : ..... : 2310

\* 2420 \*

Seqid1 : AATTTTAATCATGTGCCATATCGTTTGCGT : 2430

Seqid3 : ..... : 2340

2440 \* 2460

Seqid1 : TTTTTTGCTGGTGGCGACCAAAGTATTCGT : 2460

Seqid3 : ..... : 2370

\* 2480 \*

Seqid1 : GGATATGCACATGACAGTTTATCACCTATA : 2490

Seqid3 : ..... : 2400

WO 00/52042

PCT/EP00/01468

13/20

2500 \* 2520  
Seqid1 : TCAGATAAGGGTTATCTGACAGGCGGTCAA : 2520  
Seqid3 : ..... : 2430

\* 2540 \*  
Seqid1 : GTATTGGCGGTTGGTACAGCTGAATATAAT : 2550  
Seqid3 : ..... : 2460

2560 \* 2580  
Seqid1 : TATGAATTTATGAAAGATTTGCGTTTGGCG : 2580  
Seqid3 : ..... : 2490

\* 2600 \*  
Seqid1 : GTTTTTGGTGATATTGGTAATGCTTATGAT : 2610  
Seqid3 : ..... : 2520

2620 \* 2640  
Seqid1 : AAAGGCTTTACTAATGATACCAAATTGGT : 2640  
Seqid3 : ..... : 2550

\* 2660 \*  
Seqid1 : GCAGGTGTCGGTGTTTCGCTGGGCATCACCT : 2670  
Seqid3 : ..... : 2580

2680 \* 2700  
Seqid1 : GTCGGTCAAGTTCGTGTTGATGTGGCAACT : 2700  
Seqid3 : ..... : 2610

WO 00/52042

PCT/EP00/01468

14/20

\* 2720 \*

Seqid1 : GGTGTCAAAGAAGAGGGCAATCCCATTAAG : 2730

Seqid3 : ..... : 2640

2740 \* 2760

Seqid1 : CTGCATTTTTTTTATTGGCACACCATTTTAA : 2760

Seqid3 : ..... : 2670

WO 00/52042

PCT/EP00/01468

15/20

**Figure 2 : Alignment of the BASB081 polypeptide sequences.****Identity to SeqID No:2 is indicated by a dot. Gap is indicated by a dash.**

```

                *           20           *
Seqid2 : MSKPVL FANRSFMPVALAAYLPLMTSQALA : 30
Seqid4 : ----- : -

                40           *           60
Seqid2 : QQNNPANI INHVPAHDTAINQAKAGNPPVL : 60
Seqid4 : ..... : 30

                *           80           *
Seqid2 : LTPEQIQARLNAAGLNAKPQSQALDVVNFD : 90
Seqid4 : ..... : 60

                100           *           120
Seqid2 : DQSPISRIGE QSPPLGLDMSVIEETPLSL : 120
Seqid4 : ..... : 90

                *           140           *
Seqid2 : EELFAQESTEMGINPNDYIPEYQGEQPNSE : 150
Seqid4 : ..... : 120

                160           *           180
Seqid2 : VVVPPTLEPEKPGLIKRLYARLFNDGVNKV : 180
Seqid4 : ..... : 150
```

WO 00/52042

PCT/EP00/01468

16/20

\* 200 \*

Seqid2 : PRLKAKFYQSSQSGETSAIGSSHQKTEPYA : 210

Seqid4 : ..... : 180

220 \* 240

Seqid2 : NIKAALEDITQESAMDLNGSIPRLRQTALV : 240

Seqid4 : ..... : 210

\* 260 \*

Seqid2 : AARAVGYDIDLSIIRNSIGEVDVIIHDLG : 270

Seqid4 : ..... : 240

280 \* 300

Seqid2 : EPVYIDYRAVEVRGEGADDKAFTTVADEVP : 300

Seqid4 : ..... : 270

\* 320 \*

Seqid2 : LLIGDVFHHGKYETKKNLIENASAEHGYFD : 330

Seqid4 : ..... : 300

340 \* 360

Seqid2 : GRWLDRSVDVILPDNTADVSLIYDTGTQYR : 360

Seqid4 : ..... : 330

\* 380 \*

Seqid2 : FDEVVFFTIDPKTNQLTTDPDKLPVKRELL : 390

Seqid4 : ..... : 360



WO 00/52042

PCT/EP00/01468

17/20

400 \* 420  
Seqid2 : EQLLTVNMGEAYNLQAVRALSNDLIATRYF : 420  
Seqid4 : ..... : 390

\* 440 \*  
Seqid2 : NMVNTEIVFPEREQIQNDQVSFEQSSSSRT : 450  
Seqid4 : ..... : 420

460 \* 480  
Seqid2 : EPAQVDESTLEPVIETVELTDGILMDISPI : 480  
Seqid4 : ..... : 450

\* 500 \*  
Seqid2 : EFSASNLIQDKLNLVAAKARHLYDMPDDR : 510  
Seqid4 : ..... : 480

520 \* 540  
Seqid2 : LAINHDDGVNRSILGRISDAVSAVARAILP : 540  
Seqid4 : ..... : 510

\* 560 \*  
Seqid2 : DESENEVIDLPERTALANRKTPADVYQSKK : 570  
Seqid4 : ..... : 540

580 \* 600  
Seqid2 : VPLYVVFVASDKPRDGQIGLGWGS DTGTRLV : 600  
Seqid4 : ..... : 570

WO 00/52042

PCT/EP00/01468

18/20

\* 620 \*

Seqid2 : TKFEHNLINRDGYQAGAE LRLSEDKKGVKL : 630

Seqid4 : ..... : 600

640 \* 660

Seqid2 : YATKPLSHPLNDQLRATLGYQQEVFGHSTN : 660

Seqid4 : ..... : 630

\* 680 \*

Seqid2 : GFDLSTRTLEHEISRSIIQNGGWNRTYSLR : 690

Seqid4 : ..... : 660

700 \* 720

Seqid2 : YRLDKLKTQAPPETWQDLPVDFVNGKPSQE : 720

Seqid4 : ..... : 690

\* 740 \*

Seqid2 : ALLAGVAVHKT VADNLVNP MRGYRQ RYSLE : 750

Seqid4 : ..... : 720

760 \* 780

Seqid2 : VGSSGLVSDANMAIARAGISGVYSFGDNAY : 780

Seqid4 : ..... : 750

\* 800 \*

Seqid2 : GSNRAHQMTGGIQAGYIWSDNFNHVPYRLR : 810

Seqid4 : ..... : 780

WO 00/52042

PCT/EP00/01468

19/20

820 \* 840  
Seqid2 : FFAGGDQSIRGYAHDSLSPISDKGYLTGGQ : 840  
Seqid4 : ..... : 810

\* 860 \*  
Seqid2 : VLAVGTAEYNYEFMKDLRLAVFGDIGNAYD : 870  
Seqid4 : ..... : 840

880 \* 900  
Seqid2 : KGFTNDTKIGAGVGVRWASPVGQVRVDVAT : 900  
Seqid4 : ..... : 870

\* 919  
Seqid2 : GVKEEGNPIKLHFFIGTPF : 919  
Seqid4 : .....F : 889

WO 00/52042

PCT/EP00/01468

20/20

**Figure 3.** SDS-PAGE analysis of BASB081 expression in non induced (N.I.) or induced (I) *Escherichia coli* Top10 cells.

